

PATENT ABSTRACTS OF JAPAN

(11) Publication number : 2001-024699
 (43) Date of publication of application : 26. 01. 2001

(51) Int. Cl. H04L 12/56
 H04L 12/24
 H04L 12/26

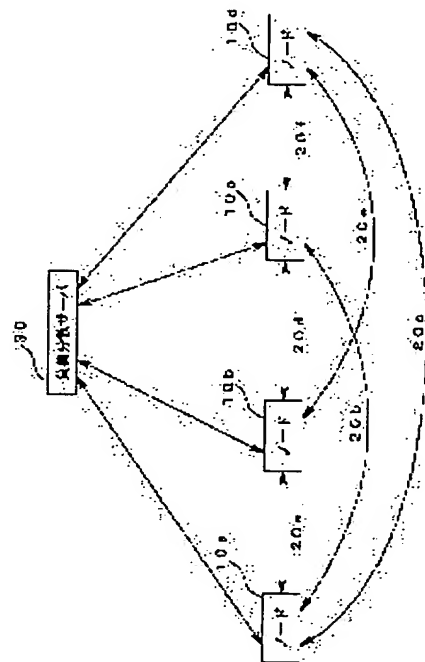
(21) Application number : 11-189665 (71) Applicant : NEC CORP
 (22) Date of filing : 02. 07. 1999 (72) Inventor : IWATA ATSUSHI

(54) NETWORK LOAD DISTRIBUTION SYSTEM

(57) Abstract:

PROBLEM TO BE SOLVED: To provide a network load distribution system, capable of optimal load distribution in consideration of whole network, and updating a routing table by deciding an optimal link metric following dynamic change related with the whole network, such as topology change and traffic fluctuations of the network.

SOLUTION: A network, in which plural nodes 10a, 10b, 10c, and 10d are mutually connected through links 20a, 20b, 20c, 20d, 20e, and 20f is provided with a load distribution server 30 for receiving network state information from the plural nodes 10a, 10b, 10c, and 10d, and deciding an optimal link metric, based on the network state information, and transmitting the optimal link metric to the plural nodes. In this case, dynamic path selection is made, based on the optimal link metric in each node 10a, 10b, 10c, and 10d.



LEGAL STATUS

[Date of request for examination] 15. 06. 2000
 [Date of sending the examiner's decision of rejection]
 [Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
 [Date of final disposal for application]
 [Patent number]
 [Date of registration]
 [Number of appeal against examiner's decision of rejection]
 [Date of requesting appeal against examiner's decision of rejection]
 [Date of extinction of right]

Copyright (C) ; 1998, 2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-24699

(P2001-24699A)

(43) 公開日 平成13年1月26日 (2001.1.26)

(51) Int.Cl.⁷

識別記号

H 0 4 L 12/56
12/24
12/26

F I

H 0 4 L 11/20
11/08

テーマコード (参考)

1 0 2 D 5 K 0 3 0
9 A 0 0 1

審査請求 有 請求項の数 11 O L (全 12 頁)

(21) 出願番号 特願平11-189665

(22) 出願日 平成11年7月2日 (1999.7.2)

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 岩田 淳

東京都港区芝五丁目7番1号 日本電気株式会社内

(74) 代理人 100093595

弁理士 松本 正夫

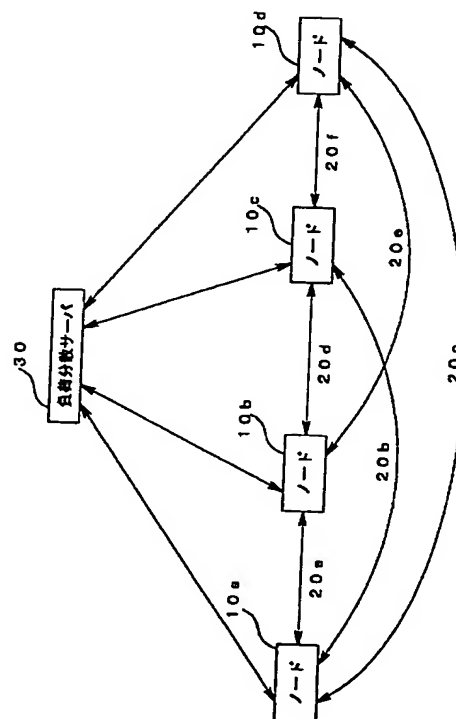
Fターム (参考) 5K030 GA01 HA10 HB06 HD03 LB07
LB08 LC01 LC06 LE03 MA01
MB09 MD07
9A001 CC03 CC06 CC07 HH32 KK56
LL02 LL09

(54) 【発明の名称】 ネットワーク負荷分散システム

(57) 【要約】

【課題】 ネットワーク全体を考慮した最適な負荷分散が可能であり、かつネットワークのトポロジー変更、トラフィック変動等のネットワーク全体に係る動的変化に追従して、最適なリンク・メトリックを判定し、ルーティング・テーブルの更新が可能なネットワーク負荷分散システムを提供する。

【解決手段】 複数のノード10a、10b、10c、10dがリンク20a、20b、20c、20d、20e、20fにより相互に接続されたネットワークにおいて、前記複数のノード10a、10b、10c、10dからネットワーク状態情報を受信し、かつ前記ネットワーク状態情報をもとに最適なリンク・メトリックを判定し、前記最適なリンク・メトリックを前記複数のノードに送信する負荷分散サーバ30を備え、各ノード10a、10b、10c、10dにおいて前記最適なリンク・メトリックに基づき動的経路選択を行なう。



【特許請求の範囲】

【請求項 1】 複数のノードがリンクにより相互に接続されたネットワークにおいて、前記複数のノードからネットワーク状態情報を受信し、かつ前記ネットワーク状態情報をもとに最適なリンク・メトリックを判定し、前記最適なリンク・メトリックを前記複数のノードに送信する負荷分散サーバを備え、各ノードにおいて前記最適なリンク・メトリックに基づき動的経路選択を行なうネットワーク負荷分散システム。

【請求項 2】 前記ネットワーク状態情報は、前記ノードの前記各隣接ノードとの接続の形態を表すネットワーク・トポロジー情報と、前記ノードから前記各隣接リンクに対する送信方向に割り当てられているリンク・メトリックの値を表すリンク・メトリック情報と、前記ノードにおけるトラフィック特性、並びにトラフィックパラメータを表すトラフィックフロー情報を含むことを特徴とする請求項 1 に記載のネットワーク負荷分散システム。

【請求項 3】 前記ネットワーク状態情報は、前記ノードの前記各隣接ノードとの接続の形態を表すネットワーク・トポロジー情報と、前記ノードから前記各隣接リンクに対する送信方向に割り当てられているリンク・メトリックの値を表すリンク・メトリック情報と、前記ノードに隣接する前記各リンクにおける前記ノードから送信方向の使用帯域、並びに使用可能最大帯域を表すリンク負荷情報と、前記ノードにおけるトラフィック特性、並びにトラフィックパラメータを表すトラフィックフロー情報を含むことを特徴とする請求項 1 に記載のネットワーク負荷分散システム。

【請求項 4】 前記負荷分散サーバは、前記ネットワーク内の前記ノード全ての経路選択の動作をエミュレーションする経路選択エミュレータを備え、前記経路選択エミュレータにより、前記ネットワーク状態情報を参照し、かつ前記経路選択エミュレータ内のリンク・メトリックの値を変化させながらエミュレーションを繰返し行なうことで、前記最適なリンク・メトリックを求めることを特徴とする請求項 1 乃至 3 に記載のネットワーク負荷分散システム。

【請求項 5】 前記負荷分散サーバは、前記経路選択エミュレータにより、前記経路選択エミュレータ内のリンク・メトリックの値を、前記ネットワーク内の前記各リンクのリンク負荷の量が等しくなるように変化させながら、前記エミュレーションを繰返し行なうことで、前記最適なリンク・メトリックを求めることを特徴とする請求項 4 に記載のネットワーク負荷分散システム。

【請求項 6】 前記負荷分散サーバは、

前記経路選択エミュレータにより、前記ネットワーク内の前記各リンクの使用帯域の量をリンク帯域の量で割った値が互いに等しくなるように、前記経路選択エミュレータ内のリンク・メトリックの値を変化させながら、前記エミュレーションを繰返し行なうことで、前記最適なリンク・メトリックを求めることを特徴とする請求項 4 に記載のネットワーク負荷分散システム。

10 【請求項 7】 前記動的経路選択は、

前記各ノードにおいて、前記最適なリンク・メトリックに基づき、前記ノードのルーティング・テーブルを更新し、さらに前記最適なリンク・メトリックをルーティング・プロトコルにより前記ノードに隣接する他の前記ノードと相互に通知し、前記隣接ノードからの情報に基づきルーティング・テーブルを更新することを特徴とする請求項 1 乃至 6 に記載のネットワーク負荷分散システム。

【請求項 8】 前記ネットワーク内の 1 つ又は複数の前記ノードは、

内部に前記負荷分散サーバを実装することを特徴とする請求項 1 乃至 7 に記載のネットワーク負荷分散システム。

【請求項 9】 前記ネットワーク内の 1 つ又は複数の前記ノードは、

内部に前記負荷分散サーバを実装し、かつ前記負荷分散サーバは、前記ルーティングプロトコルにより、前記負荷分散サーバを実装する前記ノードを、前記負荷分散サーバを実装していない他の各ノードに対し通知し、負荷分散サーバの位置を認識させることを特徴とする請求項 7 に記載のネットワーク負荷分散システム。

【請求項 10】 前記ネットワークがコネクションオリエンティッドな場合に、

前記最適なリンク・メトリックを前記隣接ノードと相互に通知する時に、既存のコネクションをも前記最適なリンク・メトリックに応じて変更することを特徴とする請求項 7 又は請求項 9 に記載のネットワーク負荷分散システム。

40 【請求項 11】 前記既存のコネクションを変更する際に、前記既存のコネクションのサービスの瞬断を発生しないように、前記既存のコネクションを残したまま新たに最適なコネクションをはり、前記既存のコネクションのサービスを新規の前記最適なコネクションに迂回させた後、前記既存のコネクションを切断することを特徴とする請求項 10 に記載のネットワーク負荷分散システム。

【発明の詳細な説明】

【0001】

50 【発明の属する技術分野】 本発明は、リンク・メトリック

を基準に最適経路を決定する動的経路選択に関し、特に、ネットワーク管理サーバによりネットワークの状態を監視し、リンク・メトリックを定期的に最適なものに更新することで、ネットワーク全体の負荷分散を実現するネットワーク負荷分散システムに関する。

【0002】

【従来の技術】ネットワークにおいてデータを送信するとき、送信先に到達可能な複数の通信経路の中から1つの経路を選択することを経路選択（またはルーティング）と言う。経路選択の方法としては、伝送時間の短縮や伝送路の使用効率の向上を目的として、最適な経路を選択するために各種の方法が使用されている。

【0003】静的経路選択は、予め各送信先に対する最適と考えられる通信経路を固定的に定めておく方式である。各送信先に対する通信経路は、ルーティング・テーブルにおいて指示され、このルーティング・テーブルを参照することで経路選択を行なう。

【0004】動的経路選択は、ルーティング・テーブルの内容を、トラフィック変動やネットワーク構成の変化に応じて最適なものに更新していく方式である。さらにこのルーティング・テーブルを、ルーティング・プロトコルを用いて各ノード間で交換することにより他のノードの情報を得て、最適な経路選択を行なう。

【0005】通信経路における中継装置の数をホップ（Hop）値といい、これによりこの通信経路による送信ノード間の距離を代表させ、実際に経由する中継装置の数が最小になる経路の中継装置の数をコストという。こうしてリンク・メトリック、つまりネットワーク上の距離を用いて最適な経路選択を行なう。

【0006】主なルーティング・プロトコルは、RIP（Routing Information Protocol）と、OSPF（Open Shortest Path First）である。

【0007】RIPは、ホップ値とコストから最適経路を決定する。しかし、コストすなわち中継装置の数が最小であっても伝送時間が最小になるとは限らない。OSPFでは、RIPのこれらの欠点を解決し、回線の輻輳状態を考慮して最適経路を決定する。

【0008】従来、この種の動的経路選択によるネットワーク負荷分散システムは、各ノードが自立的に自ノードと隣接ノードの間のリンクの負荷を測定し、負荷がある閾値を超えた時点で、事前に決められたルールに従いリンク・メトリックを適切に増減し、これに対応してルーティング・テーブルを更新し、かつこのリンク・メトリックの変更をネットワーク内の他のノードへ通知するものであった。

【0009】また、特開平05-130144号公報では、リンクの負荷状態の検出部と、リンク・メトリック変更・送信部とから構成される技術が記載されている。

【0010】

【発明が解決しようとする課題】上述したように従来のネットワーク負荷分散システムでは、以下に述べるような問題点があった。

【0011】第1に、従来のネットワーク負荷分散システムでは、各ノードの近傍のみでの局所的な負荷分散にすぎず、ネットワーク全体での最適な負荷分散ではないという問題点がある。

【0012】その理由は、各ノードが自立的に測定した自ノードと隣接ノードの間のリンクの負荷（つまり近傍リンクのみの負荷）に基づき、自ノードと隣接ノードの間のリンク・メトリックを局所的な面において適切に増減するものであったので、局所的な負荷分散にすぎず隣接ノード以外のネットワーク状況全体を考慮した負荷分散ができないからである。

【0013】第2に、前記特開平05-130144号公報に公開された従来技術では、ネットワーク全体の動的变化に応じた最適なリンク・メトリックの決定ができないという問題点がある。

【0014】その理由は、前記特開平05-130144号公報のネットワーク負荷分散システムでは、リンク負荷がある特定の閾値を超えた場合等に、閾値を超えたリンクのリンク・メトリックを一定値増減するという単純な静的な方法であるため、ネットワーク全体に係る変化であるトポロジー変更やトラフィックの分布変動等に対応する負荷分散ができないからである。

【0015】本発明の目的は、上記従来技術の欠点を解決し、ネットワークの伝送効率、信頼性、性能の向上を図るため、ネットワーク全体を考慮した最適な負荷分散が可能なネットワーク負荷分散システムを提供することである。

【0016】本発明の他の目的は、上記従来技術の欠点を解決するため、ネットワークのトポロジー変更、トラフィック変動等のネットワーク全体に係る動的变化に追従して、最適なリンク・メトリックを判定し、ルーティング・テーブルの更新が可能なネットワーク負荷分散システムを提供することである。

【0017】

【課題を解決するための手段】上記目的を達成するため本発明のネットワーク負荷分散システムは、複数のノードがリンクにより相互に接続されたネットワークにおいて、複数のノードからネットワーク状態情報を受信し、かつネットワーク状態情報をもとに最適なリンク・メトリックを判定し、最適なリンク・メトリックを複数のノードに送信する負荷分散サーバを備え、各ノードにおいて最適なリンク・メトリックに基づき動的経路選択を行なうことを特徴とする。

【0018】請求項2の本発明のネットワーク負荷分散システムにおけるネットワーク状態情報は、ノードの各隣接ノードとの接続の形態を表すネットワーク・トポロジー情報と、ノードから各隣接リンクに対する送信方向

に割り当てられているリンク・メトリックの値を表すリンク・メトリック情報と、ノードにおけるトラフィック特性、並びにトラフィックパラメータを表すトラフィックフロー情報を含むことを特徴とする。

【0019】請求項3の本発明のネットワーク負荷分散システムにおけるネットワーク状態情報は、ノードの各隣接ノードとの接続の形態を表すネットワーク・トポロジー情報と、ノードから各隣接リンクに対する送信方向に割り当てられているリンク・メトリックの値を表すリンク・メトリック情報と、ノードに隣接する各リンクにおけるノードから送信方向の使用帯域、並びに使用可能最大帯域を表すリンク負荷情報と、ノードにおけるトラフィック特性、並びにトラフィックパラメータを表すトラフィックフロー情報を含むことを特徴とする。

【0020】請求項4の本発明のネットワーク負荷分散システムにおける負荷分散サーバは、ネットワーク内のノード全ての経路選択の動作をエミュレーションする経路選択エミュレータを備え、経路選択エミュレータにより、ネットワーク状態情報を参照し、かつ経路選択エミュレータ内のリンク・メトリックの値を変化させながらエミュレーションを繰返し行なうことで、最適なリンク・メトリックを求めることを特徴とする。

【0021】請求項5の本発明のネットワーク負荷分散システムにおける負荷分散サーバは、経路選択エミュレータにより、経路選択エミュレータ内のリンク・メトリックの値を、ネットワーク内の各リンクのリンク負荷の量が等しくなるように変化させながら、エミュレーションを繰返し行なうことで、最適なリンク・メトリックを求めることを特徴とする。

【0022】請求項6の本発明のネットワーク負荷分散システムにおける負荷分散サーバは、経路選択エミュレータにより、ネットワーク内の各リンクの使用帯域の量を実リンク帯域の量で割った値が互いに等しくなるように、経路選択エミュレータ内のリンク・メトリックの値を変化させながら、エミュレーションを繰返し行なうことで、最適なリンク・メトリックを求めることを特徴とする。

【0023】請求項7の本発明のネットワーク負荷分散システムにおける動的経路選択は、各ノードにおいて、最適なリンク・メトリックに基づき、ノードのルーティング・テーブルを更新し、さらに最適なリンク・メトリックをルーティング・プロトコルによりノードに隣接する他のノードと相互に通知し、隣接ノードからの情報に基づきルーティング・テーブルを更新することを特徴とする。

【0024】請求項8の本発明のネットワーク負荷分散システムは、1つ又は複数のノードの内部に負荷分散サーバを実装することを特徴とする。

【0025】請求項9の本発明のネットワーク負荷分散システムは、ネットワーク内の1つ又は複数のノードの

内部に負荷分散サーバを実装し、かつ負荷分散サーバは、ルーティングプロトコルにより、負荷分散サーバを実装するノードを、負荷分散サーバを実装していない他の各ノードに対し通知し、負荷分散サーバの位置を認識させることを特徴とする。

【0026】請求項10の本発明のネットワーク負荷分散システムは、ネットワークがコネクションオリエンティッドな場合に、最適なリンク・メトリックを隣接ノードと相互に通知する時に、既存のコネクションをも最適なリンク・メトリックに応じて変更することを特徴とする。

【0027】請求項11の本発明のネットワーク負荷分散システムは、既存のコネクションを変更する際に、既存のコネクションのサービスの瞬断を発生しないように、既存のコネクションを残したまま新たに最適なコネクションをはり、既存のコネクションのサービスを新規の最適なコネクションに迂回させた後、既存のコネクションを切断することを特徴とする。

【0028】

【発明の実施の形態】以下、本発明の実施の形態について図面を参照して詳細に説明する。

【0029】図1は、本発明の第1の実施の形態によるネットワーク負荷分散システムの構成を示すブロック図である。

【0030】図1を参照すると、本発明の第1の実施の形態によるネットワーク負荷分散システムは、リンク20a、20b、20c、20d、20e、20fにより互いに接続された複数のノード10a、10b、10c、10dと、これら複数のノード10a、10b、10c、10dに対し接続された負荷分散サーバ30から構成される。

【0031】ノード10a、10b、10c、10dは、ルータまたは交換機から成り、各リンク20a、20b、20c、20d、20e、20fを介してデータを送受信するものである。

【0032】負荷分散サーバ30は、各ノード10a、10b、10c、10dからネットワークの状態の情報を受信し、それに応じて最適なリンク・メトリックを計算し、各ノードに通知することにより、負荷分散をはかる仕組みを持つ。

【0033】図2は、本発明の第1の実施の形態の負荷分散サーバ30の構成を示すブロック図である。

【0034】図2を参照すると、本発明の第1の実施の形態の負荷分散サーバ30は、ネットワーク状態受信部31と、リンク・メトリック判定部32と、リンク・メトリック送信部33を備える。

【0035】ネットワーク状態受信部31は、全てのノード10a、10c、10dからネットワーク状態情報を受信し、かつ、受信したネットワーク状態情報、つまりネットワーク・トポロジー情報、リンク・メトリック

情報、リンク負荷情報、トラフィックフロー情報をリンク・メトリック判定部32へ送信する。

【0036】リンク・メトリック判定部32は、受信したネットワーク状態を基に、ネットワーク全体のトポロジー、各リンク20a、20b、20c、20d、20e、20fに現在設定されているリンク・メトリック、リンク負荷、トラフィックフローのトラフィック特性、パラメータを認識し、この情報を用いて各リンク20a、20b、20c、20d、20e、20fの最適なリンク・メトリックを判定し、このリンク・メトリックをリンク・メトリック送信部33へ送信する。

【0037】リンク・メトリック判定部32は、内部に経路選択エミュレータ321を持ち、この経路選択エミュレータ321によりネットワーク上の経路選択のエミュレーションを、エミュレーションで求められるリンク負荷が適正な値となるように、リンク・メトリックの値を少しずつ修正しながらエミュレーションを繰り返し、最適なリンク・メトリックを求める。

【0038】リンク・メトリック送信部33は、リンク・メトリック判定部32で判定された最適なリンク・メトリックを各ノード10a、10b、10c、10dに対し送信する。

【0039】図3は、本発明の第1の実施の形態のノード10aの構成を示すブロック図である。

【0040】図3を参照すると、本発明の第1の実施の形態のノード10aは、ネットワーク状態検出部11と、リンク・メトリック受信部12と、リンク状態・データベース13と、ルーティング・テーブル14と、ルーティング・プロトコル15と、データ受信部16と、データ送信部17を備える。

【0041】ノード10aによるネットワーク上のデータの送受信は、データ受信部16と、データ送信部17により行なわれる。

【0042】データ受信部16は、各リンク20a、20b、20cによって接続された隣接ノード10b、10c、10dから、データを受信する。

【0043】データ送信部17は、同じく各リンク20a、20b、20cによって接続された隣接ノード10b、10c、10dに対し、そのデータの送信先とルーティング・テーブル14を参照して次に送信する隣接ノードを決定し、データを送信する。

【0044】ノード10aの、このネットワーク上のデータの送受信を負荷分散するための機能は、残るネットワーク状態検出部11と、リンク・メトリック受信部12と、リンク状態・データベース13と、ルーティング・テーブル14と、ルーティング・プロトコル15により行なわれる。

【0045】ネットワーク状態検出部11は、ノード10aのネットワークの状態を検出し負荷分散サーバ30に対し送信する。

【0046】ネットワーク状態検出部11は、ネットワーク・トポロジー検出部111と、リンク・メトリック検出部112と、リンク負荷検出部113と、トラフィックフロー検出部114を備える。

【0047】ネットワーク・トポロジー検出部111は、ノード10aが各隣接ノードと、どのリンクでどのように接続されているのかのネットワーク・トポロジー情報を検出する。つまり例えば、ノード10aとノード10bの間はリンク20aで接続されている等の検出行なう。

【0048】リンク・メトリック検出部112は、ノード10aから隣接リンク20a、20b、20cへのそれぞれに対する送信方向に、割り当てられているリンク・メトリックを検出する。方向を規定するのは、同一リンクでも送信方向と受信方向でリンク・メトリックが異なる場合があるためである。

【0049】リンク負荷検出部113は、リンク20b、20c、20dのそれぞれにおける、ノード10aから送信方向のリンク負荷を検出する。ここでリンク負荷検出部113が検出するリンク負荷の種類は、現在の使用帯域、及びリンクの使用可能最大帯域の2種類である。

【0050】トラフィックフロー検出部114は、トラフィック特性、ならびにトラフィックパラメータをフロー毎に検出する。

【0051】リンク・メトリック受信部12は、負荷分散サーバ30から送信される最適なリンク・メトリックを受信し、この最適なリンク・メトリックによりリンク状態・データベース13を更新し、この更新されたリンク状態・データベース13によりルーティング・テーブル14を再計算し更新する。ここで、ルーティング・テーブル14の再計算とは、リンク状態・データベース13における更新された最適なリンク・メトリックに基づき、ネットワーク全体で最小コストの経路を再計算し求めることである。

【0052】さらに、リンク・メトリック受信部12は、ルーティング・プロトコル15に基づき、この更新された最適なリンク・メトリックを他のノード10b、10c、10dに対し経路選択情報として送信し、かつ逆に、他のノード10b、10c、10dにおける更新された最適なリンク・メトリックが、同様に経路選択情報としてノード10aに対し送信される。

【0053】ルーティング・プロトコル15として、ルータの場合にはリンクステート型のOSPF (Open Shortest Path Fast) ルーティング・プロトコルを用いることができる。交換機の場合にも、例えばリンクステート型のPNNI (Private Network- Network Interface) ルーティング・プロトコルを用いることで同様の処理が実現できる。

【0054】次に、第1の実施の形態によるネットワーク負荷分散システムの動作について詳細に説明する。

【0055】ここでは、特にノード10aと分散サーバ30との間の動作を説明するが、他のノードとも同様である。

【0056】まず、ノード10aにおいて、ネットワーク状態検出部11を用いて、ネットワーク・トポロジー検出部111からのネットワーク・トポロジー情報と、リンク・メトリック検出部112からの現在のリンク・メトリック情報と、リンク負荷検出部113からのリンク負荷情報と、トラフィックフロー検出部114からのトラフィックフロー情報の4つのネットワーク状態を検出し負荷分散サーバ30に対し送信する。

【0057】次に、負荷分散サーバ30において、各ノード10aから送信されるネットワーク状態を、ネットワーク状態受信部31が受信し、リンク・メトリック判定部32へ送信する。

【0058】次に、リンク・メトリック判定部32では、各ノード10aから送信されるネットワーク状態情報から、現在のネットワーク全体の状態を認識し、各リンクの最適リンク・メトリックを判定し、リンク・メトリック送信部33へその値を通知する。

【0059】次に、リンク・メトリック送信部33は、ノード10aに対し、リンク20a、20b、20cのノード10bから送信方向の最適なリンク・メトリックを、送信する。

【0060】再び各ノード10aにおいて、リンク・メトリック受信部12は、最適なリンク・メトリックを負荷分散サーバ30から受信し、この最適なリンク・メトリックをリンク状態・データベース13に反映させ、それによりノード10a内のルーティング・テーブル14を再計算し更新する。ルーティング・テーブル14は、最適なリンク・メトリックを用いて、ネットワーク全体で最小コスト経路を再計算することで求める。

【0061】その後、ルーティング・プロトコル15により、更新されたリンク・メトリック情報を隣接ノード10b、10c、10dへ経路選択情報として配布する。一方、他のノード10b、10c、10dも、同様に更新されたリンク・メトリック情報をノード10aへ経路選択情報として配布してくるので、その情報をルーティング・プロトコル15が受信すると、その情報をまずリンク状態・データベース13へ格納する。

【0062】その後、リンク状態・データベース13の更新に伴い、再び最小コスト経路を再計算してルーティング・テーブル14を更新する。

【0063】以上の手続きにより、すべてのノード10a、10b、10c、10dが、自分の新規のリンク・メトリック情報を更新するとともに、その情報を他のノードにも通知し合うことにより、それぞれのノードが新規のリンク・メトリック情報に従ったルーティング・テ

ーブル14で経路選択を行なう。

【0064】つまり、負荷分散サーバ30内のリンク・メトリック判定部32の計算した通りに、実際にトラフィックが負荷分散されることとなる。

【0065】負荷分散サーバ30と各ノード10a、10b、10c、10dの以上の処理は定期的に行われ、時間が経過する毎にトポロジー、リンク・メトリック、リンク負荷、トラフィックフローの変化に動的に対応し、常にトラフィック負荷分散が図られる。

【0066】定期的な実行の周期は、ネットワークの変動のスピードと、負荷分散に対する要求条件にしたがって決める。

【0067】図4は、本発明の第1の実施の形態のリンク・メトリック判定部32の、最適なリンク・メトリックの判定処理を説明するためのフローチャートである。

【0068】この判定処理においては、各変数を以下のように定める。

【0069】全ノードの数を「N」、ネットワーク内のノードを、「ノードi」、「ノードj」($i, j \leq N$)、ノードiとノードj間の全経路の数を

「 K_{ij} 」、隣接するノードiからノードjへの、現在のリンク・メトリックの値を「 A_{ij} 」、隣接するノードiからノードjへの、求める最適化されたリンク・メトリックの値を「 B_{ij} 」、隣接するノードiからノードjへの、現在のリンク負荷を「 C_{ij} 」、ネットワーク内のノードiからノードjへの、経路k ($k \leq K_{ij}$) によるトラフィックフローを「 T_{ijk} 」と表す。

【0070】まず、リンク・メトリック判定部32は、ネットワーク状態受信部31を介して、ネットワークの全ノード(ノード1～ノードN)からネットワーク状態情報を受信する(ステップ401)。

【0071】このネットワーク状態情報とは、ネットワーク・トポロジー情報と、リンク・メトリック情報{ A_{ij} : ノードi, jは隣接}と、リンク負荷情報{ C_{ij} : ノードi, jは隣接}と、トラフィックフロー情報{ T_{ijk} : $i \neq j, k \leq K_{ij}$ }との全ての情報である。

【0072】次に、受信したこれらのネットワーク状態情報を基に、最適な新規リンク・メトリック B_{ij} を以下の手順で計算する。

【0073】リンク・メトリック判定部32の内部の、経路選択エミュレータ321のカウンタを"0"にリセットする(ステップ402)。

【0074】ネットワーク状態情報を内部の経路選択エミュレータ321に初期情報として設定する(ステップ403)。ここで、変数リンク・メトリック値

{ A_{ij} }、リンク負荷{ C_{ij} }、トラフィックフロー{ T_{ijk} }は、経路選択エミュレータ321内の計算用のメモリにそれぞれ{ a_{ij} }、{ c_{ij} }、{ T

$i j k$ }として格納する。

【0075】経路選択エミュレータ321により、実際の全ノード1～ノードNが行うルーティング・テーブル14の計算、並びにこのルーティング・テーブル14によるデータの経路選択と同一の処理をエミュレートする。

【0076】経路選択エミュレータ321は、ネットワーク・トポロジー情報とリンク・メトリック $\{a_{ij}\}$ とから、まず最小コスト経路を計算し、擬似的に経路選択エミュレータ321内の全ノードのルーティング・テーブル14を更新する(ステップ404)。

【0077】ここでもし複数の経路が同一の最小コスト経路である場合、すべての経路をエミュレータ内ルーティング・テーブル14に登録し、それら複数の経路をトラフィックが負荷分散するように経路制御を行う。

【0078】その後、トラフィックフロー $\{T_{ijk}\}$ をこのエミュレータ内のノードへ仮想的に流して、それによってトラフィックが具体的にどの経路を流れていくかを計算する(ステップ405)。

【0079】すべてのトラフィックフロー $\{T_{ijk}\}$ に対して、新しい経路が決まると、その時点で全ノード1～ノードN間のリンクに予想されるリンク負荷 $\{c_{ij} : \text{ノード } i, j \text{ は隣接}\}$ が求められる(ステップ406)。

【0080】手順225では、これら求められたリンク負荷 $\{c_{ij}\}$ の中で最大の負荷を持つリンク、ノード x からノード y ($x \leq N, y \leq N$ 、ノード x, y は隣接)を選択する。そして、現在のリンク・メトリック $\{a_{ij}\}$ の値に対して、このリンク $\{x, y\}$ に対応するリンク・メトリックの値 a_{xy} のみに正の数 δ を足し、これを新しいリンク・メトリック $\{a_{ij}\}$ として定める(ステップ407)。

【0081】カウンタを1増加させ(ステップ408)、もしカウンタがある閾値M以上ならば、最適なリンク・メトリックを求める反復計算がある一定の回数を超えたものとみなして、エミュレーションの繰返しを終了し(ステップ409)、その時点で得られたリンク・メトリック $\{a_{ij}\}$ を、新規リンク・メトリック $\{B_{ij}\}$ として、リンク・メトリック送信部33へ送信する(ステップ411)。

【0082】また、リンク負荷 $\{c_{ij}\}$ の各リンクでの値がすべて同じとなる場合にも、既に最適なリンク・メトリックが求められたものとしてエミュレーションの繰返しを終了し(ステップ410)、その時点で得られたリンク・メトリック $\{a_{ij}\}$ を、新規リンク・メトリック $\{B_{ij}\}$ として、リンク・メトリック送信部33へ送信する(ステップ411)。

【0083】このいずれの場合でもない時には、再びステップ404に戻り、新しいリンク・メトリック $\{a_{ij}\}$ に基づきエミュレーションを繰返す。

【0084】以上により得られた最適なリンク・メトリックは、リンク・メトリック送信部33により各ノード1～Nに送信される。

【0085】また、以上説明されたリンク・メトリック判定部32の判定処理においては、初期情報であるリンク負荷情報 $\{C_{ij}\}$ を使用していないが、この情報は、負荷の量を負荷分散を行なうかどうかの判断として用いる場合、つまり、ある負荷を超えるまで負荷分散を行なわない等の場合に用いることができる。さらに、経路選択エミュレータ321が1回目のエミュレーションにより求めたリンク負荷 $\{c_{ij}\}$ の値と、大きな違いがないかを参照することにより、ネットワーク状態情報を正しく受信しているかのチェックに用いることができる。しかしこのため、リンク負荷情報 $\{C_{ij}\}$ は、ネットワーク状態情報の中に含めないものとしても良い。

【0086】同様に、初期情報であるトラフィックフロー情報 $\{T_{ijk}\}$ は、トラフィックフローの量の情報として、経路別の情報が必要ではなく、全ての二つのノード間の情報が得られれば良いが、これも、経路選択エミュレータ321が1回目のエミュレーションにより求めたトラフィックフロー情報 $\{t_{ijk}\}$ の値と、大きな違いがないかを参照することにより、ネットワーク状態情報を正しく受信しているかのチェックに用いることができる。しかしこのため、トラフィックフロー情報は、経路別に集めなくとも良い。

【0087】以上のように、本実施の形態では、負荷分散サーバ30がネットワーク内の全ノード10a, 10b, 10c, 10dから、ネットワーク状態を集散的に集め、その情報に基づいて、最適なリンク・メトリックを判定し、動的経路選択をするため、ネットワーク全体での最適な負荷分散を行うことができる。

【0088】また、本実施の形態では、定期的にネットワーク状態を取得して、最適なリンク・メトリックを判定するため、ネットワークのトポロジー変更や、トラフィック変動等のネットワーク全体に関する動的な変化に素早く対応して負荷分散を行うことができる。

【0089】次に、本発明の第2の実施の形態について、図面を参照して詳細に説明する。

【0090】図5は、本発明の第2の実施の形態によるネットワーク負荷分散システムの構成を示すブロック図である。

【0091】図5を参照すると、本発明の第2の実施の形態によるネットワーク負荷分散システムの、図1における第1の実施の形態によるネットワーク負荷分散システムとの違いは、負荷分散サーバ30aを、データ伝送を行なう各ノードから独立に設置せずに、ハードウェア的に任意の1つのノード40aの内部に負荷分散サーバ30aの機能を実装したことである。

【0092】本実施の形態のネットワーク負荷分散システムの機能的には第1の実施の形態と同じ動作を行う。

負荷分散サーバ 30a を実装するノード 40a の情報を、ルーティング・プロトコル 15 によって交換させることにより（たとえば OSPF では Opaque LSA を用いて交換可能である）、負荷分散サーバ 30a を実装しない全てのノード 40b, 40c, 40d が自動的に負荷分散サーバ 30a を実装するノード 40a の位置を認識し、負荷分散サーバ 30a とのやりとりを実現する。

【0093】また、負荷分散サーバ 30a を実装するノード 40a は、1 つでなくとも良く複数のノード内に実装させることができる。この複数のノード内に実装させる場合には、各ノードに実装する負荷分散サーバ 30a はそれぞれ同等のものであっても、機能を分割して各ノードに実装させてもよい。

【0094】以上のように、本実施の形態では、ノード 40a の内部に負荷分散サーバ 30a の機能を実装するために、負荷分散サーバ 30a をデータ伝送を行なう各ノードから独立に設置する必要や、また、負荷分散サーバ 30a と各ノード 40a, 40b, 40c, 40d との通信のために特別に回線を繋ぐ必要がない。このため、第 1 の実施の形態の効果に加えて、より簡潔な構成で設置も容易なネットワーク負荷分散システムを実現できる。

【0095】本発明の第 3 の実施の形態について、図面を参照して詳細に説明する。

【0096】図 6 は、本発明の第 3 の実施の形態によるネットワーク負荷分散システムの構成を示すブロック図である。

【0097】図 6 を参照すると、本実施の形態によるネットワーク負荷分散システムは、ATM 交換機 50a, 50b, 50c, 50d を使用したコネクショレス型のネットワークで実施されている。

【0098】第 1 の実施の形態によるネットワーク負荷分散システムでは、主にコネクショレス型の経路選択の場合について説明した。本発明を、ATM のように、各データの通信経路を伝送開始前にあらかじめ設定する方式であるコネクショ型のネットワークで実施する場合には、以下のように一部制御手順を追加することが望ましい。

【0099】負荷分散サーバ 30b が、ネットワーク状態情報に対応し、最適な新規リンク・メトリックを計算し、各 ATM 交換機 50a, 50b, 50c, 50d に通知が終わると、第 1 の実施の形態を直接適応した場合には、ATM 交換機 50a, 50b, 50c, 50d は伝送中のセルについては通信経路の設定を変更せず、新規に伝送要求のあったセルに対してのみ、更新された最適なリンク・メトリックを用いて通信経路を設定し負荷分散を行う。

【0100】つまり、既存のコネクションには何も変化を加えない手順である。したがって、第 3 の実施の形態

としては、ATM 交換機 50a, 50b, 50c, 50d が最適な新規リンク・メトリックを受け取った時点で、既存のコネクションをすべて新規リンク・メトリックに応じてすべて経路選択を行わせる場合がある。

【0101】この実施の形態の場合、経路選択を行っている間、コネクションが切断され、既存のコネクション上のサービスが瞬断する可能性があるため、既存の VCC (Virtual Channel Connection) を残したまま、まず最適な VCC を設定し、既存のサービスを新規設定した VCC 上に迂回が終了した後に、既存の VCC を切断するなどのサービスの瞬断を防ぐ方式を併用する。

【0102】以上のように、本実施の形態では、第 1 の実施の形態の効果に加えて、コネクショ型のネットワークにおいて負荷分散を行なう時に、既に転送中の通信経路が設定されているデータに対しても最新の最適な経路に設定を変更することにより、性能、効率の良いネットワーク負荷分散システムを実現できる。

【0103】以上好ましい実施の形態及び実施例をあげて本発明を説明したが、本発明は必ずしも上記実施の形態及び実施例に限定されるものではなく、その技術的思想の範囲内において様々に変形して実施することができる。

【0104】例えば、負荷分散の方針は図 4 に説明された上述の方法と同様にして、ネットワーク管理者の意図に従い、さまざまな方法をとることが可能である。

【0105】負荷分散を行なうかどうかの判断として、ある負荷を超えるまではまったく負荷分散はおこなわず、ある負荷の閾値を超えた瞬間から負荷分散を行う方法や、たとえ負荷が低くても常に負荷分散を行う方法が、負荷分散の初期情報であるネットワーク状態情報中のリンク負荷情報を参照することにより可能である。

【0106】また、負荷の占める相対的な割合を均等にする負荷分散を行う方法や、負荷による残余帯域を均等にする負荷分散を行う方法が、リンク負荷情報をそれぞれ（使用帯域／実リンク帯域）つまり使用帯域の量を実リンク帯域の量で割った値として設定したり、または負荷による残余帯域の値として設定することで、図 4 のフローチャートの流れにより同様に負荷分散が可能である。

【0107】また、第 3 の実施の形態と第 2 の実施の形態を組み合わせることも可能である。つまり第 2 の実施の形態の方法により、第 3 の実施の形態の任意の 1 つの ATM 交換機 50a の内部に負荷分散サーバ 30a の機能を持たせることが可能である。

【0108】

【発明の効果】第 1 に、ネットワーク全体での最適な負荷分散を行うことができる。

【0109】その理由は、負荷分散サーバがネットワーク内の全ノードから、ネットワーク状態を集中的に集め、その情報に基づいて、最適な負荷分散をはかるため

の最適なリンク・メトリックを判定し、それを各ノードに配布することと、さらに、その最適なリンク・メトリックがルーティング・プロトコルによってすべてのノードに行き渡り、各ノードのルーティング・テーブルがそのリンク・メトリックに基づき更新されるためである。

【0110】第2に、ネットワークのトポロジ変更や、トラフィック変動等のネットワーク全体に関する動的な変化に対応して、最適なリンク・メトリックを決定し、ルーティング・テーブルを変更できる。

【0111】その理由は、定期的にネットワーク状態を

【図面の簡単な説明】

【図1】 本発明の第1の実施の形態によるネットワーク負荷分散システムの構成を示すブロック図である。

【図2】 本発明の第1の実施の形態の負荷分散サーバの構成を示すブロック図である。

【図3】 本発明の第1の実施の形態のノードの構成を示すブロック図である。

【図4】 本発明の第1の実施の形態のリンク・メトリック判定部の、最適なリンク・メトリックの判定処理を説明するためのフローチャートである。

【図5】 本発明の第2の実施の形態によるネットワーク負荷分散システムの構成を示すブロック図である。

【図6】 本発明の第3の実施の形態によるネットワーク負荷分散システムの構成を示すブロック図である。

【符号の説明】

10a、10b、10c、10d、40a、40b、40c、40d ノード

50a、50b、50c、50d ATM交換機

20a、20b、20c、20d、20e、20f リンク

11 ネットワーク状態検出部

111 ネットワーク・トポロジ検出部

112 リンク・メトリック検出部

113 リンク負荷検出部

114 トラフィックフロー検出部

12 リンク・メトリック受信部

13 リンク状態・データベース

14 ルーティング・テーブル

15 ルーティング・プロトコル

16 データ受信部

17 データ送信部

30、30a、30b 負荷分散サーバ

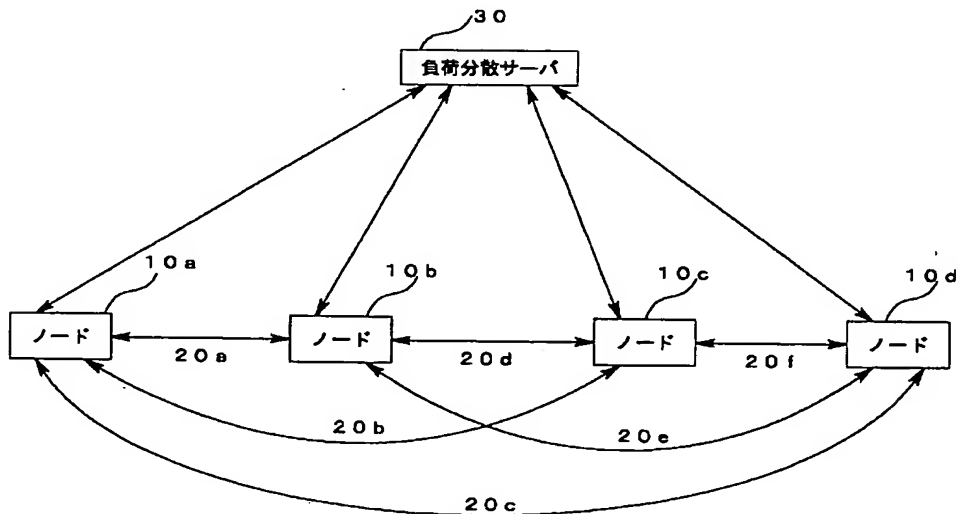
31 ネットワーク状態受信部

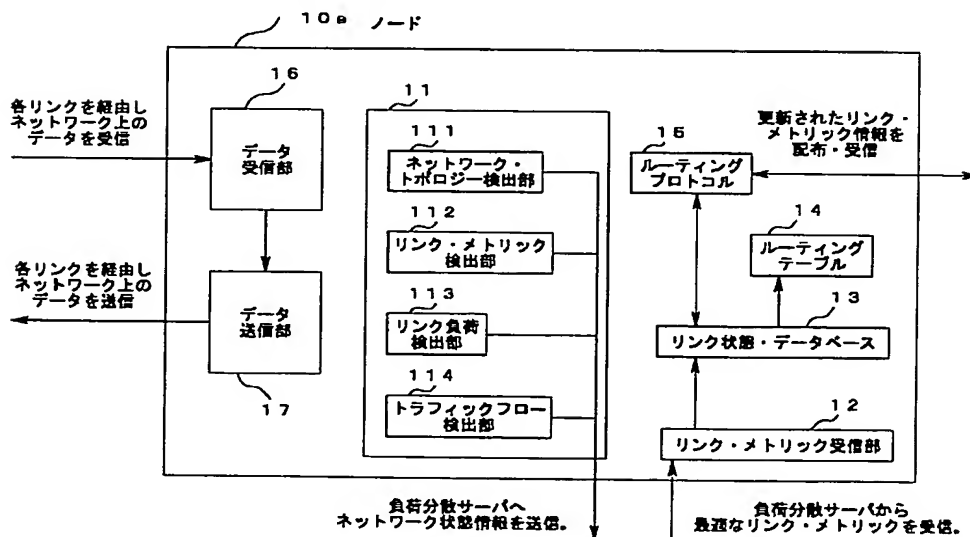
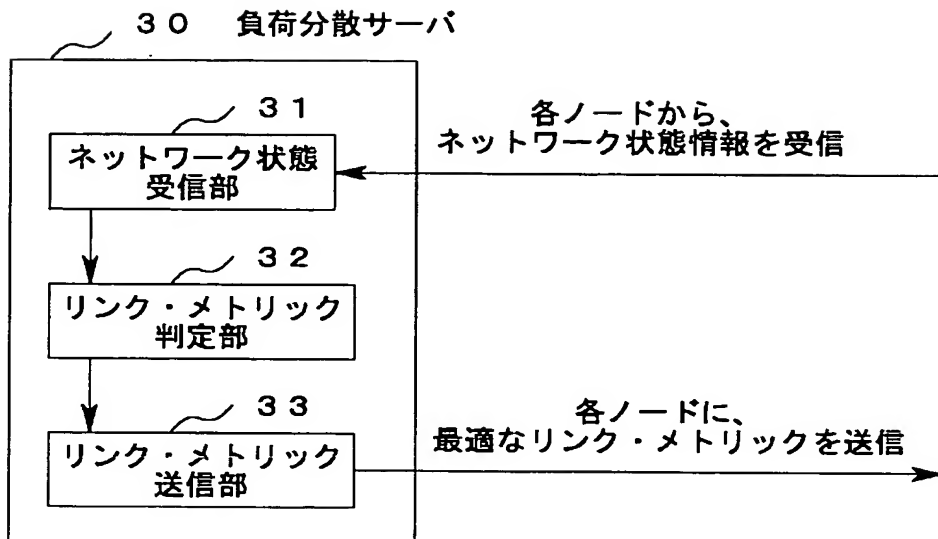
32 リンク・メトリック判定部

321 経路選択エミュレータ

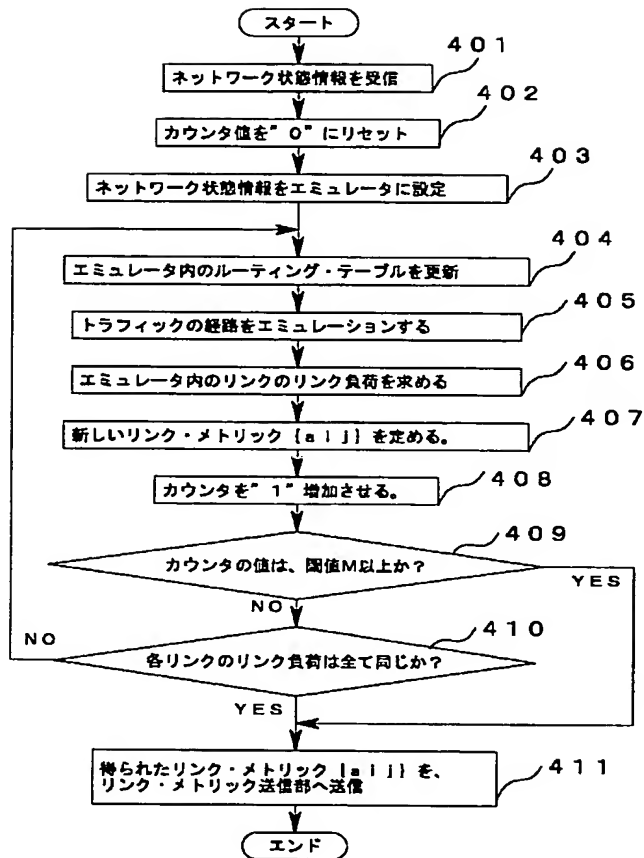
33 リンク・メトリック送信部

【図1】

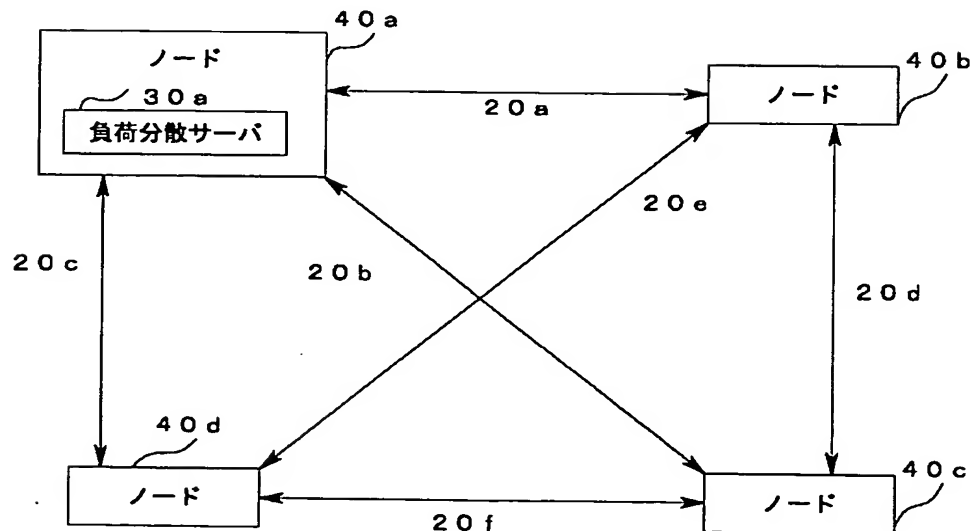


vi

【図 4】



【図 5】



【図6】

